

SEQUENCE LISTING

(1) GENERAL INFORMATION

- (i) APPLICANT: Lal, Preeti
Corley, Neil C.
Yue, Henry
- (ii) TITLE OF THE INVENTION: HUMAN E1-LIKE PROTEIN
- (iii) NUMBER OF SEQUENCES: 4
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: Incyte Pharmaceuticals, Inc.
 - (B) STREET: 3174 Porter Drive
 - (C) CITY: Palo Alto
 - (D) STATE: CA
 - (E) COUNTRY: USA
 - (F) ZIP: 94304
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Diskette
 - (B) COMPUTER: IBM Compatible
 - (C) OPERATING SYSTEM: DOS
 - (D) SOFTWARE: FastSEQ for Windows Version 2.0
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER: To Be Assigned
 - (B) FILING DATE: Herewith
 - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER:
 - (B) FILING DATE:
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Billings, Lucy J.
 - (B) REGISTRATION NUMBER: 36,749
 - (C) REFERENCE/DOCKET NUMBER: PF-0487 US
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: 650-855-0555
 - (B) TELEFAX: 650-845-4166
 - (C) TELEX:

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 363 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (vii) IMMEDIATE SOURCE:
 - (A) LIBRARY: UTRSNOT11
 - (B) CLONE: 2546462
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

PF-0487-2 DIV

```

Met Ala Val Asp Gly Gly Cys Gly Asp Thr Gly Asp Trp Glu Gly Arg
 1      5      10      15
Trp Asn His Val Lys Lys Phe Leu Glu Arg Ser Gly Pro Phe Thr His
 20      25      30
Pro Asp Phe Glu Pro Ser Thr Glu Ser Leu Gln Phe Leu Leu Asp Thr
 35      40      45
Cys Lys Val Leu Val Ile Gly Ala Gly Gly Leu Gly Cys Glu Leu Leu
 50      55      60
Lys Asn Leu Ala Leu Ser Gly Phe Arg Gln Ile His Val Ile Asp Met
 65      70      75      80
Asp Thr Ile Asp Val Ser Asn Leu Asn Arg Gln Phe Leu Phe Arg Pro
 85      90      95
Lys Asp Ile Gly Arg Pro Lys Ala Glu Val Ala Ala Glu Phe Leu Asn
100      105      110
Asp Arg Val Pro Asn Cys Asn Val Val Pro His Phe Asn Lys Ile Gln
115      120      125
Asp Phe Asn Asp Thr Phe Tyr Arg Gln Phe His Ile Ile Val Cys Gly
130      135      140
Leu Asp Ser Ile Ile Ala Arg Arg Trp Ile Asn Gly Met Leu Ile Ser
145      150      155      160
Leu Leu Asn Tyr Glu Asp Gly Val Leu Asp Pro Ser Ser Ile Val Pro
165      170      175
Leu Ile Asp Gly Gly Thr Glu Gly Phe Lys Gly Asn Ala Arg Val Ile
180      185      190
Leu Pro Gly Met Thr Ala Cys Ile Glu Cys Thr Leu Glu Leu Tyr Pro
195      200      205
Pro Gln Val Asn Phe Pro Met Cys Thr Ile Ala Ser Met Pro Arg Leu
210      215      220
Pro Glu His Cys Ile Glu Tyr Val Arg Met Leu Gln Trp Pro Lys Glu
225      230      235      240
Gln Pro Phe Gly Glu Gly Val Pro Leu Asp Gly Asp Asp Pro Glu His
245      250      255
Ile Gln Trp Ile Phe Gln Lys Ser Leu Glu Arg Ala Ser Gln Tyr Asn
260      265      270
Ile Arg Gly Val Thr Tyr Arg Leu Thr Gln Gly Val Val Lys Arg Ile
275      280      285
Ile Pro Ala Val Ala Ser Thr Asn Ala Val Ile Ala Ala Val Cys Ala
290      295      300
Thr Glu Val Phe Lys Ile Ala Thr Ser Ala Tyr Ile Pro Leu Asn Asn
305      310      315      320
Tyr Leu Val Phe Asn Asp Val Asp Gly Leu Tyr Thr Tyr Thr Phe Glu
325      330      335
Ala Glu Arg Lys Val Ser Ser Ile Lys Asn Thr Phe Leu Ile Met His
340      345      350
Ile Leu Ile Phe Lys Tyr Tyr Trp Leu Glu Ile
355      360

```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2084 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: UTRSNOT11
- (B) CLONE: 2546462

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```

AACAAATATGG CGGATGGCGA GGAGCGGAGA AGAAAAGAAG GAGAATAGAG GAGCTGCTGG      60
CTGAGAAAAAT GGCTGTTGAT GGTGGGTGTG GGGACACTGG AGACTGGGAA GGTCGCTGGA      120
ACCATGTAAA GAAGTTCCTC GAGCGATCTG GACCCTTCAC ACACCTGAT TTCGAACCGA      180
GCACTGAATC TCTCCAGTTC TTGTTAGATA CATGTAAAGT TCTAGTCATT GGAGCTGGCG      240
GCTTAGGATG TGAGCTCCTG AAAAATCTGG CCTTGTCTGG TTTTAGACAG ATTCATGTTA      300
TAGATATGGA CACTATAGAT GTTTCCAATC TAAATAGGCA GTTTTATTTT AGGCCTAAAG      360
ATATTGGAAG ACCTAAGGCT GAAGTTGCTG CAGAATTTCT AAATGACAGA GTTCCTAATT      420
GCAATGTAGT TCCACATTTT AACAAGATTG AAGATTTTAA CGACACTTTC TATCGACAAT      480
TTCATATTAT TGTATGTGGA CTGGACTCTA TCATCGCCAG AAGATGGATA AATGGCATGC      540
TGATATCTCT TCTAAATTAT GAAGATGGTG TCTTAGATCC AAGCTCCATT GTCCCTTTGA      600
TAGACGGGGG GACAGAAGGT TTAAAGGAA ATGCCCCGGT GATTCTGCCT GGAATGACTG      660
CTTGTATCGA ATGCACGCTG GAACTTTATC CACCACAGGT TAATTTTCCC ATGTGCACCA      720
TTGCATCTAT GCCCAGGCTA CCAGAACACT GTATTGAGTA TGTAAAGGATG TTGCAGTGGC      780
CTAAGGAGCA GCCTTTTGGG GAAGGGGTTC CATTAGATGG AGATGATCCT GAACATATAC      840
AATGGATTTT CCAAAAAATCC CTAGAGAGAG CATCACAATA TAATATTAGG GGTGTTACGT      900
ATAGGCTCAC TCAAGGGGTA GTAAAAAGAA TCATTCTCTG AGTAGCTTCC ACAAATGCAG      960
TCATTGCAGC TGTGTGTGCC ACTGAGGTTT TTAAAATAGC CACAAGTGCA TACATTCCCT      1020
TGAATAATTA CTTGGTGTTT AATGATGTAG ATGGGCTGTA TACATACACA TTTGAAGCAG      1080
AAAGAAAGGT TAGTAGTATT AAGAACACAT TTTTGATCAT GCATATTTTG ATTTTAAAT      1140
ATTATTGGTT AGAAATTTGA ACAAAGTCAC CCATACATTT TCTAACTTCC AGAACTCTAC      1200
TTATTATATA TCTTTTGCTT TATAGCCTGA AATAACTCTA TAGCGAAGTA ATTTACAAGA      1260
AATGGTCTAT TATGAAAAGC AGGCTTTAAA GCATAAAAT TTTTTTATAG GAAATATGCA      1320
TGATTATAAA ACAACCTGAT TTTTATTTTA TTGTTCATAA AAGAGACTAA TATTGGTGCA      1380
TGTGCTGCTG TAATTTGTTG TGTATTATGT GTGTAGGAAA ACTGCCCAGC TTGTAGCCAG      1440
CTTCCTCAA ATATTAGTT TTCTCCATCA GCTAAACTAC AGGAGGTTTT GGATTATCTA      1500
ACCAATAGTG CTTCTCTGTA AGTATTGTAG ATTTTGTGTA TGTGTAAAA ATCATTTTGT      1560
TGATTTTGA AACCTTAAAA AAATTATCTT TTGATAAAAA TTATGTTTGA TACTTCTCTC      1620
TCATCATAAT CTTTAGGCAA ATGAAATCTC CAGCCATCAC AGCCACCTA GAGGGAAAAA      1680
ATAGAACACT TTAATTACAG GTTATCAATG TGTATTTTAA ATTTTTTTC AAAAAATTATA      1740
TCAAGTTTTT TTTTACTTTA ATGTGTCTTA CATTAAAGTA ATTTTGTTTT CTAGTCGGTA      1800
ACCTCTATTG AAGAACGAAC AAGGCCAAAT CTCTCCAAAA CATTGAAAGG TATTTTACAT      1860
AAGGGTATTT ACTAATCATT TTCTTTCTTT TCTCTCTTTT TGGTGAAAGT AATCAGTGCT      1920
TGTTCTAGAT TTCCTCTTAA TGCCTTGTAT ATGGTCAGGT AATAATTACT TACAACCTTA      1980
GACATATTAA TAGAATTAAT TGCTCTTTTA GTAGGATATT TAAAATCTCC AAGGAATCAA      2040
TATTTACTTT GATTAAAGAG GATTGGNTTT TGATGTTTTN CTAG      2084

```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 437 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: GenBank
- (B) CLONE: 1055197

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```

Met Val Ser Val Asp Pro Leu Ala Thr Glu Arg Trp Arg Ser Ile Arg
1          5          10          15
Arg Leu Thr Asp Arg Asp Ser Ala Tyr Lys Val Pro Trp Phe Val Pro
20          25          30
Gly Pro Glu Asn Phe Glu Ala Leu Gln Asn Thr Lys Ile Leu Val Ile
35          40          45
Gly Ala Gly Gly Leu Gly Cys Glu Leu Leu Lys Asn Leu Ala Leu Ser
50          55          60
Gly Phe Arg Thr Ile Glu Val Ile Asp Met Asp Thr Ile Asp Val Ser
65          70          75          80

```

```

Asn Leu Asn Arg Gln Phe Leu Phe Arg Glu Ser Asp Val Gly Lys Ser
      85          90          95
Lys Ala Glu Val Ala Ala Ala Phe Val Gln Gln Arg Val Val Gly Cys
      100          105          110
Gln Asn Tyr Phe Asn Phe Ile Ser Ile Phe Arg His Asn Cys Arg Ile
      115          120          125
Glu Asp Lys Gly Gln Glu Phe Tyr Arg Lys Phe Ser Ile Ile Ile Cys
      130          135          140
Gly Leu Asp Ser Ile Pro Ala Arg Arg Trp Ile Asn Gly Met Leu Cys
      145          150          155          160
Asp Leu Val Leu Glu Met Ala Asp Gly Lys Pro Asp Glu Asn Thr Ile
      165          170          175
Ile Pro Met Ile Asp Gly Gly Thr Glu Gly Phe Lys Gly Asn Ala Arg
      180          185          190
Val Ile Tyr Pro Lys Phe Thr Ala Cys Ile Asp Cys Thr Leu Asp Leu
      195          200          205
Tyr Pro Pro Gln Val Asn Phe Pro Leu Cys Thr Ile Ala His Thr Pro
      210          215          220
Arg Leu Pro Glu His Cys Ile Glu Tyr Ile Lys Val Val Val Trp Pro
      225          230          235          240
Glu Glu Lys Pro Phe Glu Gly Val Ser Leu Asp Ala Asp Asp Pro Ile
      245          250          255
His Val Glu Trp Val Leu Glu Arg Ala Ser Leu Arg Ala Glu Lys Tyr
      260          265          270
Asn Ile Arg Gly Val Asp Arg Arg Leu Thr Ser Gly Val Leu Lys Arg
      275          280          285
Ile Ile Pro Ala Val Ala Ser Thr Asn Ala Val Ile Ala Ala Ser Cys
      290          295          300
Ala Leu Glu Ala Leu Lys Leu Ala Thr Asn Ile Ala Lys Pro Ile Asp
      305          310          315          320
Asn Tyr Leu Asn Phe Thr Gln Ile His Gly Ala Tyr Thr Ser Val Val
      325          330          335
Ser Met Met Lys Asp Asp Asn Cys Leu Thr Cys Ser Gly Gly Arg Leu
      340          345          350
Pro Phe Glu Val Ser Pro Ser Ser Thr Leu Glu Ser Leu Ile Ile Arg
      355          360          365
Leu Ser Glu Arg Phe His Leu Lys His Pro Thr Leu Ala Thr Ser Thr
      370          375          380
Arg Lys Leu Tyr Cys Ile Ser Ser Phe Met Pro Gln Phe Glu Gln Glu
      385          390          395          400
Ser Lys Glu Asn Leu His Thr Ser Met Lys Asp Leu Val Ser Asp Gly
      405          410          415
Glu Glu Ile Leu Val Ser Asp Glu Ala Leu Ser Arg Ala Leu Thr Leu
      420          425          430
Arg Ile Gln Leu Ile
      435

```

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 636 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: GenBank
- (B) CLONE: 793879

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

PF-0487-2 DIV

Met	Pro	Arg	Glu	Thr	Ser	Leu	Val	Thr	Ile	Ile	Gly	Glu	Asp	Ser	Tyr	1	5	10	15
Lys	Lys	Leu	Arg	Ser	Ser	Arg	Cys	Leu	Leu	Val	Gly	Ala	Gly	Gly	Ile	20	25	30	
Gly	Ser	Glu	Leu	Leu	Lys	Asp	Ile	Ile	Leu	Met	Glu	Phe	Gly	Glu	Ile	35	40	45	
His	Ile	Val	Asp	Leu	Asp	Thr	Ile	Asp	Leu	Ser	Asn	Leu	Asn	Arg	Gln	50	55	60	
Phe	Leu	Phe	Arg	Gln	Lys	Asp	Ile	Lys	Gln	Pro	Lys	Ser	Thr	Thr	Ala	65	70	75	80
Val	Lys	Ala	Val	Gln	His	Phe	Asn	Asn	Ser	Lys	Leu	Val	Pro	Tyr	Gln	85	90	95	
Gly	Asn	Val	Met	Asp	Ile	Ser	Thr	Phe	Pro	Leu	His	Trp	Phe	Glu	Gln	100	105	110	
Phe	Asp	Ile	Ile	Phe	Asn	Ala	Leu	Asp	Asn	Leu	Ala	Ala	Arg	Arg	Tyr	115	120	125	
Val	Asn	Lys	Ile	Ser	Gln	Phe	Leu	Ser	Leu	Pro	Leu	Ile	Glu	Ser	Gly	130	135	140	
Thr	Ala	Gly	Phe	Asp	Gly	Tyr	Met	Gln	Pro	Ile	Ile	Pro	Gly	Lys	Thr	145	150	155	160
Glu	Cys	Phe	Glu	Cys	Thr	Lys	Lys	Glu	Thr	Pro	Lys	Thr	Phe	Pro	Val	165	170	175	
Cys	Thr	Ile	Arg	Ser	Thr	Pro	Ser	Gln	Pro	Ile	His	Cys	Ile	Val	Trp	180	185	190	
Ala	Lys	Asn	Phe	Leu	Phe	Asn	Gln	Leu	Phe	Ala	Ser	Glu	Thr	Ser	Gly	195	200	205	
Asn	Glu	Asp	Asp	Asn	Asn	Gln	Asp	Trp	Gly	Thr	Asp	Asp	Ala	Glu	Glu	210	215	220	
Ile	Lys	Arg	Ile	Lys	Gln	Glu	Thr	Asn	Glu	Leu	Tyr	Glu	Leu	Gln	Lys	225	230	235	240
Ile	Ile	Ile	Ser	Arg	Asp	Ala	Ser	Arg	Ile	Pro	Glu	Ile	Leu	Asn	Lys	245	250	255	
Leu	Phe	Ile	Gln	Asp	Ile	Asn	Lys	Leu	Leu	Ala	Ile	Glu	Asn	Leu	Trp	260	265	270	
Lys	Thr	Arg	Thr	Lys	Pro	Val	Pro	Leu	Ser	Asp	Ser	Gln	Ile	Asn	Thr	275	280	285	
Pro	Thr	Lys	Thr	Ala	Gln	Ser	Ala	Ser	Asn	Ser	Val	Gly	Thr	Ile	Gln	290	295	300	
Glu	Gln	Ile	Ser	Asn	Phe	Ile	Asn	Ile	Thr	Gln	Lys	Leu	Met	Asp	Arg	305	310	315	320
Tyr	Pro	Lys	Glu	Gln	Asn	His	Ile	Glu	Phe	Asp	Lys	Asp	Asp	Ala	Asp	325	330	335	
Thr	Leu	Glu	Phe	Val	Ala	Thr	Ala	Ala	Asn	Ile	Arg	Ser	His	Ile	Phe	340	345	350	
Asn	Ile	Pro	Met	Lys	Ser	Val	Phe	Asp	Ile	Lys	Gln	Ile	Ala	Gly	Asn	355	360	365	
Ile	Ile	Pro	Ala	Ile	Ala	Thr	Asn	Ala	Ile	Val	Ala	Gly	Ala	Ser		370	375	380	
Ser	Leu	Ile	Ser	Leu	Arg	Val	Leu	Asn	Leu	Leu	Lys	Tyr	Ala	Pro	Thr	385	390	395	400
Thr	Lys	Tyr	Thr	Asp	Leu	Asn	Met	Ala	Phe	Thr	Ala	Lys	Ala	Ser	Asn	405	410	415	
Leu	Ser	Gln	Asn	Arg	Tyr	Leu	Ser	Asn	Pro	Lys	Leu	Ala	Pro	Pro	Asn	420	425	430	
Lys	Asn	Cys	Pro	Val	Cys	Ser	Lys	Val	Cys	Arg	Gly	Val	Ile	Lys	Leu	435	440	445	
Ser	Ser	Asp	Cys	Leu	Asn	Lys	Met	Lys	Leu	Ser	Asp	Phe	Val	Val	Leu	450	455	460	
Ile	Arg	Glu	Lys	Tyr	Ser	Tyr	Pro	Gln	Asp	Ile	Ser	Leu	Leu	Asp	Ala	465	470	475	480

PF-0487-2 DIV

Ser	Asn	Gln	Arg	Leu	Leu	Phe	Asp	Tyr	Asp	Phe	Glu	Asp	Leu	Asn	Asp	
				485					490					495		
Arg	Thr	Leu	Ser	Glu	Ile	Asn	Leu	Gly	Asn	Gly	Ser	Ile	Ile	Leu	Phe	
			500					505					510			
Ser	Asp	Glu	Glu	Gly	Asp	Thr	Met	Ile	Arg	Lys	Ala	Ile	Glu	Leu	Phe	
		515					520						525			
Leu	Asp	Val	Asp	Asp	Glu	Leu	Pro	Cys	Asn	Thr	Cys	Ser	Leu	Pro	Asp	
	530					535						540				
Val	Glu	Val	Pro	Leu	Ile	Lys	Ala	Asn	Asn	Ser	Pro	Ser	Lys	Asn	Glu	
545					550					555					560	
Glu	Glu	Glu	Lys	Asn	Glu	Lys	Gly	Ala	Asp	Val	Val	Ala	Thr	Thr	Asn	
				565					570					575		
Ser	His	Gly	Lys	Asp	Gly	Ile	Val	Ile	Leu	Asp	Asp	Asp	Glu	Gly	Glu	
			580					585					590			
Ile	Thr	Ile	Asp	Ala	Glu	Pro	Ile	Asn	Gly	Ser	Lys	Lys	Arg	Pro	Val	
		595					600					605				
Asp	Thr	Glu	Ile	Ser	Glu	Ala	Pro	Ser	Asn	Lys	Arg	Thr	Lys	Leu	Val	
	610					615					620					
Asn	Glu	Pro	Thr	Asn	Ser	Asp	Ile	Val	Glu	Leu	Asp					
625					630					635						